

(19) 日本国特許庁(JP)

## (12) 公開特許公報(A)

(11) 特許出願公開番号

特開2004-184769

(P2004-184769A)

(43) 公開日 平成16年7月2日(2004.7.2)

(51) Int.CI.<sup>7</sup>  
G10G 3/04F1  
G10G 3/04テーマコード(参考)  
5D082

		審査請求 未請求 請求項の数 10 O.L. (全 20 頁)
(21) 出願番号	特願2002-352865 (P2002-352865)	(71) 出願人 000005016 パイオニア株式会社 東京都目黒区目黒1丁目4番1号
(22) 出願日	平成14年12月4日 (2002.12.4)	(74) 代理人 100079119 弁理士 藤村 元彦 (72) 発明者 義山 真一 埼玉県鶴ヶ島市富士見6丁目1番1号 パイオニア株式会社総合研究所内 Fターム(参考) 5D082 BB01 BB21

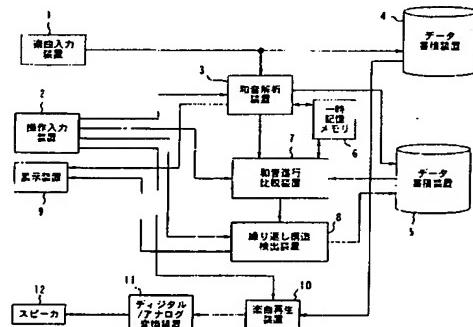
## (54) 【発明の名称】 楽曲構造検出装置及び方法

## (57) 【要約】

【課題】 繰り返し部分を含む楽曲の構造を簡単な構成で適切に検出することができる楽曲構造検出装置及び方法を提供する。

【解決手段】 楽曲の和音の時系列変化を示す和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成し、部分楽曲データ各々と和音進行楽曲データとを和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して複数の楽曲毎の類似度を算出し、部分楽曲データ毎に算出された類似度各々に応じて類似度が所定値より高いピーク値となった和音進行楽曲データ中の和音の位置を検出し、和音進行楽曲データ中の和音の位置毎に部分楽曲データ全てについて類似度が所定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出力を生成する。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

楽曲の和音の時系列変化を示す和音進行楽曲データに応じてその楽曲の構造を検出する楽曲構造検出装置であって、

前記和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成する部分楽曲データ生成手段と、

前記部分楽曲データ各々と前記和音進行楽曲データとを前記和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して前記複数の楽曲毎の類似度を算出する比較手段と、

前記部分楽曲データ毎に前記比較手段によって算出された類似度各々に応じて類似度が所定値より高いピーク値となった前記和音進行楽曲データ中の和音の位置を検出する和音位置検出手段と、

前記和音進行楽曲データ中の和音の位置毎に前記部分楽曲データ全てについて前記類似度が前記所定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する出力手段と、を備えたことを特徴とする楽曲構造検出装置。

10

**【請求項 2】**

前記比較手段は、前記部分楽曲データ各々と前記和音進行楽曲データとを前記和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とに加えて和音変化前後の和音の時間的長さの比について比較して前記複数の楽曲毎の類似度を算出することを特徴とする請求項 1 記載の楽曲構造検出装置。

20

**【請求項 3】**

前記比較手段は、前記部分楽曲データ各々と前記和音進行楽曲データとを時間的に前後に飛び越して比較することを特徴とする請求項 1 記載の楽曲構造検出装置。

**【請求項 4】**

前記比較手段は、前記部分楽曲データ各々が示す和音変化後の和音と前記和音進行楽曲データが示す和音変化後の和音とが関係調にあるときその双方の和音変化後の和音を同一の和音とみなすことを特徴とする請求項 1 記載の楽曲構造検出装置。

30

**【請求項 5】**

前記部分楽曲データ各々と前記和音進行楽曲データとは和音変化時点毎に 2 つの和音を第 1 及び第 2 和音候補として有し、

前記比較手段は、前記部分楽曲データ各々の第 1 及び第 2 和音候補と前記和音進行楽曲データの第 1 及び第 2 和音候補とを相互に比較することを特徴とする請求項 1 記載の楽曲構造検出装置。

40

**【請求項 6】**

楽曲を示す入力オーディオ信号を所定の時間毎に周波数成分の大きさを示す周波数信号に変換する周波数変換手段と、

前記周波数変換手段によって得られた周波数信号から平均律の各音に対応した周波数成分を前記所定の時間毎に抽出する成分抽出手段と、

前記成分抽出手段によって抽出された各音に対応した周波数成分のうちのレベル合計が大となる 3 つの周波数成分の組によって各々形成される 2 つの和音を前記第 1 及び第 2 和音候補として検出する和音候補検出手段と、

前記和音候補検出手段によって繰り返し検出された前記第 1 及び第 2 和音候補各々の列を平滑化処理して前記蓄積手段に記憶させるべき前記和音進行楽曲データを生成する平滑化手段と、を備えたことを特徴とする請求項 5 記載の楽曲構造検出装置。

**【請求項 7】**

前記比較手段は、前記和音進行楽曲データの終了部分に前記所定数の仮想和音だけの仮想データを付加して前記部分楽曲データ各々との比較に用いることを特徴とする請求項 1 記載の楽曲構造検出装置。

**【請求項 8】**

50

前記出力手段は、前記和音進行楽曲データ中の和音の位置毎の算出回数が最高回数の部分の楽曲音を再生して出力する請求項1記載の楽曲構造検出装置。

**【請求項9】**

楽曲の和音の時系列変化を示す和音進行楽曲データに応じてその楽曲の構造を検出する楽曲構造検出方法であって、

前記和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成する部分楽曲データ生成ステップと、

前記部分楽曲データ各々と前記和音進行楽曲データとを前記和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して前記複数の楽曲毎の類似度を算出する比較ステップと、

前記部分楽曲データ毎に前記比較ステップにおいて算出された類似度各々に応じて類似度が所定値より高いピーク値となった前記和音進行楽曲データ中の和音の位置を検出する和音位置検出ステップと、

前記和音進行楽曲データ中の和音の位置毎に前記部分楽曲データ全てについて前記類似度が前記所定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する出力ステップと、を備えたことを特徴とする楽曲構造検出方法。

**【請求項10】**

楽曲の和音の時系列変化を示す和音進行楽曲データに応じてその楽曲の構造を検出する方法を実行するコンピュータ読み取可能なプログラムであって、

前記和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成する部分楽曲データ生成ステップと、

前記部分楽曲データ各々と前記和音進行楽曲データとを前記和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して前記複数の楽曲毎の類似度を算出する比較ステップと、

前記部分楽曲データ毎に前記比較ステップにおいて算出された類似度各々に応じて類似度が所定値より高いピーク値となった前記和音進行楽曲データ中の和音の位置を検出する和音位置検出ステップと、

前記和音進行楽曲データ中の和音の位置毎に前記部分楽曲データ全てについて前記類似度が前記所定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する出力ステップと、を備えたことを特徴とするプログラム。

**【発明の詳細な説明】**

**【0001】**

**【発明が属する技術分野】**

本発明は、楽曲の和音の時系列変化を示すデータに応じてその楽曲の構造を検出する楽曲構造検出装置及び方法に関する。

**【0002】**

**【従来の技術】**

ポピュラー音楽の楽曲においては、フレーズ（楽句）がイントロ、Aメロ、Bメロ、サビのように表現され、Aメロ、Bメロやサビの各フレーズは楽曲中で通常何回か繰り返される。楽曲中のいわゆる盛り上がり部分であるサビのフレーズは、ラジオやテレビの音楽番組やコマーシャルで最も演奏されるところである。このようなフレーズは放送する際にその楽曲音を実際に聴取して判断することが一般的である。

**【0003】**

**【発明が解決しようとする課題】**

ところで、楽曲のサビ等のフレーズがどのように繰り返されているかなどの楽曲全体の構造を知ることができれば、サビ部分に限らず、他の繰り返しフレーズ部分を容易に選択的に演奏することができる。しかしながら、従来、楽曲全体の構造を自動的に検出する装置はなく、上記したように利用者が実際に聴取して判断するしかなかった。

10

20

30

40

50

**【0004】**

そこで、本発明が解決しようとする課題には、上記の問題点が一例として挙げられ、繰り返し部分を含む楽曲の構造を簡単な構成で適切に検出することができる楽曲構造検出装置及び方法を提供することが本発明の目的である。

**【0005】****【課題を解決するための手段】**

本発明の楽曲構造検出装置は、楽曲の和音の時系列変化を示す和音進行楽曲データに応じてその楽曲の構造を検出する楽曲構造検出装置であって、前記和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成する部分楽曲データ生成手段と、前記部分楽曲データ各々と前記和音進行楽曲データとを前記和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して前記複数の楽曲毎の類似度を算出する比較手段と、前記部分楽曲データ毎に前記比較手段によって算出された類似度各々に応じて類似度が所定値より高いピーク値となった前記和音進行楽曲データ中の和音の位置を検出する和音位置検出手段と、前記和音進行楽曲データ中の和音の位置毎に前記部分楽曲データ全てについて前記類似度が前記所定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する出力手段と、を備えたことを特徴としている。10

**【0006】**

本発明の楽曲構造検出装置は、楽曲の和音の時系列変化を示す和音進行楽曲データに応じてその楽曲の構造を検出する楽曲構造検出方法であって、前記和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成する部分楽曲データ生成ステップと、前記部分楽曲データ各々と前記和音進行楽曲データとを前記和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して前記複数の楽曲毎の類似度を算出する比較ステップと、前記部分楽曲データ毎に前記比較ステップにおいて算出された類似度各々に応じて類似度が所定値より高いピーク値となった前記和音進行楽曲データ中の和音の位置を検出する和音位置検出ステップと、前記和音進行楽曲データ中の和音の位置毎に前記部分楽曲データ全てについて前記類似度が前記所定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する出力ステップと、を備えたことを特徴としている。20

**【0007】**

本発明のプログラムは、楽曲の和音の時系列変化を示す和音進行楽曲データに応じてその楽曲の構造を検出する方法を実行するコンピュータ読み取り可能なプログラムであって、前記和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成する部分楽曲データ生成ステップと、前記部分楽曲データ各々と前記和音進行楽曲データとを前記和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して前記複数の楽曲毎の類似度を算出する比較ステップと、前記部分楽曲データ毎に前記比較ステップにおいて算出された類似度各々に応じて類似度が所定値より高いピーク値となった前記和音進行楽曲データ中の和音の位置を検出する和音位置検出ステップと、前記和音進行楽曲データ中の和音の位置毎に前記部分楽曲データ全てについて前記類似度が前記所定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する出力ステップと、を備えたことを特徴としている。30

**【0008】****【発明の実施の形態】**

以下、本発明の実施例を図面を参照しつつ詳細に説明する。

図1は本発明を適用した楽曲処理システムを示している。この楽曲処理システムは、楽曲入力装置1、操作入力装置2、和音解析装置3、データ蓄積装置4、5、一時記憶メモリ6、和音進行比較装置7、繰り返し構造検出装置8、表示装置9、楽曲再生装置10、デジタル／アナログ変換装置11及びスピーカ12を備えている。40

## 【0009】

楽曲入力装置1は和音解析装置3及びデータ蓄積装置5に接続され、デジタル化されたオーディオ信号（例えば、PCMデータ）を再生する装置であり、例えば、CDプレーヤである。操作入力装置2は本システムに対してユーザが操作してデータや指令を入力するための装置である。操作入力装置2の出力は和音解析装置3、和音進行比較装置7、繰り返し構造検出装置8及び楽曲再生装置10に接続されている。データ蓄積装置4には楽曲入力装置1から供給された楽曲データ（PCMデータ）がファイルとして記憶される。

## 【0010】

和音解析装置3は、供給された楽曲データの和音を後述する和音解析動作によって解析する。一時記憶メモリ6には和音解析装置3によって解析された楽曲データの各和音が第1及び第2和音候補として一時的に記憶される。データ蓄積装置5には和音解析装置3によって解析されて和音進行楽曲データが楽曲毎にファイルとして記憶される。10

## 【0011】

和音進行比較装置7は、データ蓄積装置5に記憶された和音進行楽曲データとその和音進行楽曲データ中の一部分である部分楽曲データとを後述するように比較して類似度を算出する。繰り返し構造検出装置8は和音進行比較装置7の比較結果を用いて楽曲の繰り返し部分を検出する。

表示装置9には繰り返し構造検出装置8によって検出された繰り返し部分を含む楽曲構造が表示される。

## 【0012】

楽曲再生装置10は、繰り返し構造検出装置8によって検出された繰り返し部分の楽曲データをデータ蓄積装置4から読み出して再生し、デジタルオーディオ信号として順次出力する。デジタル／アナログ変換装置11は楽曲再生装置10によって再生されたデジタルオーディオ信号をアナログオーディオ信号に変換してスピーカ12に供給する。20

## 【0013】

和音解析装置3、和音進行比較装置7、繰り返し構造検出装置8及び楽曲再生装置10各々は操作入力装置2からの指令に応じて動作する。

次に、かかる構成の楽曲処理システムの動作について説明する。

ここでは楽曲入力装置1から出力される楽曲音を示すデジタルオーディオ信号が和音解析装置3に供給されたとする。30

## 【0014】

上記した和音解析動作としては前処理、本処理及び後処理がある。和音解析装置3は前処理として周波数誤差検出動作を行う。

周波数誤差検出動作においては、図2に示すように、時間変数T及び帯域データF(N)が0に初期化され、更に変数Nの範囲が-3～3の如く初期設定される（ステップS1）。入力デジタル信号に対してフーリエ変換によって周波数変換を0.2秒間隔で行うことによって周波数情報f(T)が得られる（ステップS2）。

## 【0015】

今回のf(T)、前回のf(T-1)及び前々回のf(T-2)を用いて移動平均処理が行われる（ステップS3）。この移動平均処理では、0.6秒以内では和音が変化することが少ないという仮定で過去2回分の周波数情報が用いられる。移動平均処理は次式によって演算される。40

$$f(T) = (f(T) + f(T-1) / 2.0 + f(T-2) / 3.0) / 3.0 \dots \dots (1)$$

ステップS3の実行後、変数Nが-3に設定され（ステップS4）、その変数Nは4より小であるか否かが判別される（ステップS5）。N < 4の場合には、移動平均処理後の周波数情報f(T)から周波数成分f1(T)～f5(T)が各々抽出される（ステップS6～S10）。周波数成分f1(T)～f5(T)は、(110.0 + 2 × N)Hzを基本周波数とした5オクターブ分の平均律の12音のものである。12音はA, A#, B, C, C#, D, D#, E, F, F#, G, G#である。図3はA音を1.0とした場合の12音及び1オクターブ高いA音各々の周波数比を示している。ステップS6のf1(T)はA音を(1.1

$0.0 + 2 \times N$ )Hz とし、ステップ S 7 の  $f_2(T)$  は A 音を  $2 \times (110.0 + 2 \times N)$ Hz とし、ステップ S 8 の  $f_3(T)$  は A 音を  $4 \times (110.0 + 2 \times N)$ Hz とし、ステップ S 9 の  $f_4(T)$  は A 音を  $8 \times (110.0 + 2 \times N)$ Hz とし、ステップ S 10 の  $f_5(T)$  は A 音を  $16 \times (110.0 + 2 \times N)$ Hz としている。

#### 【0016】

ステップ S 6 ~ S 10 の実行後、周波数成分  $f_1(T) \sim f_5(T)$  は 1 オクターブ分の帯域データ  $F'(T)$  に変換される（ステップ S 11）。帯域データ  $F'(T)$  は、  

$$F'(T) = f_1(T) \times 5 + f_2(T) \times 4 + f_3(T) \times 3 + f_4(T) \times 2 + f_5(T) \dots \dots \quad (2)$$
の如く表される。すなわち、周波数成分  $f_1(T) \sim f_5(T)$  各々は個別に重み付けされた後、加算される。1 オクターブの帯域データ  $F'(T)$  は、帯域データ  $F(N)$  に加算される  
（ステップ S 12）。その後、変数 N には 1 が加算され（ステップ S 13）、そして、ステップ S 5 が再度実行される。  
10

#### 【0017】

ステップ S 6 ~ S 13 の動作は、ステップ S 5において N が 4 より小、すなわち -3 ~ +3 の範囲であると判断される限り繰り返される。これによって音成分  $F(N)$  は -3 ~ +3 の範囲の音程誤差を含む 1 オクターブ分の周波数成分となる。

ステップ S 5において  $N \geq 4$  と判別された場合には、変数 T が所定値 M より小であるか否かが判別される（ステップ S 14）。 $T < M$  の場合には、変数 T に 1 が加算され（ステップ S 15）、ステップ S 2 が再度実行される。M 回分の周波数変換による周波数情報  $f(T)$  に対して変数 N 毎の帯域データ  $F(N)$  が算出される。  
20

#### 【0018】

ステップ S 14において  $T \geq M$  と判別された場合には、変数 N 每の 1 オクターブ分の帯域データ  $F(N)$  のうちの各周波数成分の総和が最大値となる  $F(N)$  が検出され、その検出  $F(N)$  の N が誤差値 X として設定される（ステップ S 16）。この前処理によって誤差値 X を求めることによってオーケストラの演奏音等の楽曲音全体の音程が平均律と一定の差をもっている場合に、それを補償して後述の和音解析の本処理を行うことができる。

#### 【0019】

前処理の周波数誤差検出動作が終了すると、和音解析動作の本処理が行われる。なお、誤差値 X が既に分かっている場合やその誤差を無視できる場合には、前処理は省略しても良い。本処理では楽曲全部について和音解析が行われるために楽曲の最初の部分から入力デイジタル信号は和音解析装置 3 に供給されるとする。  
30

#### 【0020】

本処理においては、図 4 に示すように、入力デイジタル信号に対してフーリエ変換によって周波数変換を 0.2 秒間隔で行うことによって周波数情報  $f(T)$  が得られる（ステップ S 21）。このステップ S 21 が周波数変換手段に対応する。そして、今回の  $f(T)$ 、前回の  $f(T-1)$  及び前々回の  $f(T-2)$  を用いて移動平均処理が行われる（ステップ S 22）。ステップ S 21 及び S 22 は上記したステップ S 2 及び S 3 と同様に実行される。  
40

#### 【0021】

ステップ S 22 の実行後、移動平均処理後の周波数情報  $f(T)$  から周波数成分  $f_1(T) \sim f_5(T)$  が各々抽出される（ステップ S 23 ~ S 27）。上記したステップ S 6 ~ S 10 と同様に、周波数成分  $f_1(T) \sim f_5(T)$  は、 $(110.0 + 2 \times N)$ Hz を基本周波数とした 5 オクターブ分の平均律の 12 音 A, A#, B, C, C#, D, D#, E, F, F#, G, G# である。ステップ S 23 の  $f_1(T)$  は A 音を  $(110.0 + 2 \times N)$ Hz とし、ステップ S 24 の  $f_2(T)$  は A 音を  $2 \times (110.0 + 2 \times N)$ Hz とし、ステップ S 25 の  $f_3(T)$  は A 音を  $4 \times (110.0 + 2 \times N)$ Hz とし、ステップ S 26 の  $f_4(T)$  は A 音を  $8 \times (110.0 + 2 \times N)$ Hz とし、ステップ S 27 の  $f_5(T)$  は A 音を  $16 \times (110.0 + 2 \times N)$ Hz としている。ここで、N はステップ S 16 で設定された X である。  
40

#### 【0022】

ステップ S 23 ~ S 27 の実行後、周波数成分  $f_1(T) \sim f_5(T)$  は 1 オクターブ分の帯域データ  $F'(T)$  に変換される（ステップ S 28）。このステップ S 28 も上記のステッ  
50

・ プ S 1 1 と同様に式(2)を用いて実行される。帯域データ F' (T)は各音成分を含むことになる。ステップ S 2 3 ~ S 2 8 が成分抽出手段に相当する。

#### 【0023】

ステップ S 2 8 の実行後、帯域データ F' (T)中の各音成分のうちの強度レベルが大きいものから 6 音が候補として選択され(ステップ S 2 9)、その 6 音候補から 2 つの和音 M 1, M 2 が作成される(ステップ S 3 0)。候補の 6 音のうちから 1 つの音を根音(ルート)として 3 音からなる和音が作成される。すなわち<sub>6</sub>C<sub>3</sub>通りの組み合わせの和音が考慮される。各和音を構成する 3 音のレベルが加算され、その加算結果の値が最大となった和音が第 1 和音候補 M 1 とされ、加算結果の値が 2 番目に大きい和音が第 2 和音候補 M 2 とされる。

10

#### 【0024】

帯域データ F' (T)の各音成分が図 5 に示すように 12 音に対する強度レベルを示す場合には、ステップ S 2 9 では A, E, C, G, B, D の 6 音が選択される。その 6 音 A, E, C, G, B, D のうちの 3 音から作成される 3 和音は、(A, C, E)からなる和音 A m、(音 C, E, G)からなる和音 C、(音 E, B, G)からなる和音 E m、(音 G, B, D)からなる和音 G、……の如くである。和音 A m(音 A, C, E)の合計強度レベルは 12、和音 C(音 C, E, G)の合計強度レベルは 9、和音 E m(音 E, B, G)の合計強度レベルは 7、和音 G(音 G, B, D)の合計強度レベルは 4 である。よって、ステップ S 3 0 では和音 A m の合計強度レベル 12 が最大となるので、第 1 和音候補 M 1 として和音 A m が設定され、和音 C の合計強度レベル 7 が 2 番目に大きいので、第 2 和音候補 M 2 として和音 C が設定される。

20

#### 【0025】

また、帯域データ F' (T)の各音成分が図 6 に示すように 12 音に対する強度レベルを示す場合には、ステップ S 2 9 では C, G, A, E, B, D の 6 音が選択される。その 6 音 C, G, A, E, B, D のうちの 3 音から作成される 3 和音は、(音 C, E, G)からなる和音 C、(A, C, E)からなる和音 A m、(音 E, B, G)からなる和音 E m、(音 G, B, D)からなる和音 G、……の如くである。和音 C(音 C, E, G)の合計強度レベルは 11、和音 A m(音 A, C, E)の合計強度レベルは 10、和音 E m(音 E, B, G)の合計強度レベルは 7、和音 G(音 G, B, D)の合計強度レベルは 6 である。よって、ステップ S 3 0 では和音 C の合計強度レベル 11 が最大となるので、第 1 和音候補 M 1 として和音 C が設定され、和音 A m の合計強度レベル 10 が 2 番目に大きいので、第 2 和音候補 M 2 として和音 A m が設定される。

30

#### 【0026】

和音を構成する音は 3 音に限らず、セブンスやディミニッシュセブンス等の 4 音もある。4 音からなる和音に対しては図 7 に示すように 3 音からなる 2 つ以上の和音に分類されている。よって、4 音からなる和音に対しても 3 音からなる和音と同様に、帯域データ F' (T)の各音成分の強度レベルに応じて 2 つの和音候補を設定することができる。

40

#### 【0027】

ステップ S 3 0 の実行後、ステップ S 3 0 において設定された和音候補数があるか否かが判別される(ステップ S 3 1)。ステップ S 3 0 では少なくとも 3 つの音を選択するだけの強度レベルに差がない場合には和音候補が全く設定されないことになるので、ステップ S 3 1 の判別が行われる。和音候補数 > 0 である場合には、更に、その和音候補数が 1 より大であるか否かが判別される(ステップ S 3 2)。

#### 【0028】

ステップ S 3 1 において和音候補数 = 0 と判別された場合には前回 T - 1 (約 0.2 秒前)の本処理において設定された和音候補 M 1, M 2 が今回の和音候補 M 1, M 2 として設定される(ステップ S 3 3)。ステップ S 3 2 において和音候補数 = 1 と判別された場合には今回のステップ S 3 0 の実行では第 1 和音候補 M 1 だけが設定されたので、第 2 和音候補 M 2 は第 1 和音候補 M 1 と同一の和音に設定される(ステップ S 3 4)。ステップ S 2 9 ~ S 3 4 が和音候補検出手段に相当する。

50

## 【0029】

ステップS32において和音候補数>1と判別された場合には今回のステップS30の実行では第1及び第2和音候補M1, M2の両方が設定されたので、時刻、第1及び第2和音候補M1, M2が一時記憶メモリ6に記憶される(ステップS35)。一時記憶メモリ6には図8に示すように時刻、第1和音候補M1、第2和音候補M2が1組となって記憶される。時刻は0.2秒毎に増加するTで表される本処理実行回数である。そのTの順に第1及び第2和音候補M1, M2が記憶される。

## 【0030】

具体的には、一時記憶メモリ6に各和音候補を図8に示したように1バイトで記憶させるために、基本音(根音)とその属性との組み合わせが用いられる。基本音には平均律の12音が用いられ、属性にはメジャー{4, 3}、マイナー{3, 4}、セブンス候補{4, 6}及びディミニッシュセブンス(dim7)候補{3, 3}の和音の種類が用いられる。{}内は半音を1とした場合の3音の差である。本来、セブンス候補は{4, 3, 3}及びディミニッシュセブンス(dim7)候補{3, 3, 3}であるが、3音で示すために上記のように表示している。

10

## 【0031】

基本音の12音は図9(a)に示すように16ビット(16進表記)で表され、属性の和音の種類は同様に図9(b)に示すように16ビット(16進表記)で表される。その基本音の下位4ビットと属性の下位4ビットがその順に連結されて図9(c)に示すように8ビット(1バイト)として和音候補として用いられる。ステップS35はステップS33又はS34を実行した場合にもその直後に実行される。

20

## 【0032】

ステップS35の実行後、楽曲が終了したか否かが判別される(ステップS36)。例えば、デジタルオーディオ信号の入力がなくなった場合、或いは操作入力装置2からの楽曲の終了を示す操作入力があった場合には楽曲が終了したと判断される。これによって本処理が終了する。

楽曲の終了が判断されるまでは変数Tに1が加算され(ステップS37)、ステップS21が再度実行される。ステップS21は上記したように0.2秒間隔で実行され、前回の実行時から0.2秒が経過して再度実行される。

30

## 【0033】

後処理においては、図10に示すように、一時記憶メモリ6から全ての第1及び第2和音候補がM1(0)～M1(R)及びM2(0)～M2(R)として読み出される(ステップS41)。0は開始時刻であり、開始時刻の第1及び第2和音候補がM1(0)及びM2(0)である。Rは最終時刻であり、最終時刻の第1及び第2和音候補がM1(R)及びM2(R)である。読み出された第1和音候補M1(0)～M1(R)及び第2和音候補M2(0)～M2(R)について平滑化が行われる(ステップS42)。この平滑化は和音の変化時点とは関係なく0.2秒間隔で和音候補を検出したことにより和音候補に含まれるノイズによる誤差を除去するために行われる。平滑化の具体的方法としては、3つの連続する第1和音候補M1(t-1), M1(t), M1(t+1)についてM1(t-1)≠M1(t)かつM1(t)≠M1(t+1)の関係が成立するか否かが判別され、その関係が成立する場合には、M1(t+1)にM1(t)は等しくされる。この判別は第1和音候補毎に行われる。第2和音候補についても同様の方法により平滑化は行われる。なお、M1(t+1)にM1(t)を等しくするのではなく、逆に、M1(t+1)をM1(t)に等しくしても良い。

40

## 【0034】

平滑化後、第1及び第2和音候補の入れ替え処理が行われる(ステップS43)。一般的に0.6秒のような短い期間には和音が変化する可能性は低い。しかしながら、信号入力段の周波数特性及び信号入力時のノイズによって帯域データF'(T)中の各音成分の周波数が変動することによって第1及び第2和音候補が0.6秒以内に入れ替わることが起ることがあり、これに対処するためにステップS43は行われる。第1及び第2和音候補が入れ替えの具体的方法としては、5つの連続する第1和音候補M1(t-2), M1(t-

50

1), M<sub>1</sub>(t), M<sub>1</sub>(t+1), M<sub>1</sub>(t+2)及びそれに対応する5つの連続する第2和音候補M<sub>2</sub>(t-2), M<sub>2</sub>(t-1), M<sub>2</sub>(t), M<sub>2</sub>(t+1), M<sub>2</sub>(t+2)についての次の如き判別が実行される。すなわち、M<sub>1</sub>(t-2)=M<sub>1</sub>(t+2), M<sub>2</sub>(t-2)=M<sub>2</sub>(t+2), M<sub>1</sub>(t-1)=M<sub>1</sub>(t)=M<sub>1</sub>(t+1)=M<sub>2</sub>(t-2)及びM<sub>2</sub>(t-1)=M<sub>2</sub>(t)=M<sub>2</sub>(t+1)=M<sub>1</sub>(t-2)の関係が成立するか否かが判別される。この関係が成立する場合には、M<sub>1</sub>(t-1)=M<sub>1</sub>(t)=M<sub>1</sub>(t+1)=M<sub>1</sub>(t-2)及びM<sub>2</sub>(t-1)=M<sub>2</sub>(t)=M<sub>2</sub>(t+1)=M<sub>2</sub>(t-2)が定められ、M<sub>1</sub>(t-2)とM<sub>2</sub>(t-2)と間で和音の入れ替えが行われる。なお、M<sub>1</sub>(t-2)とM<sub>2</sub>(t-2)との間で和音の入れ替えに代えてM<sub>1</sub>(t+2)とM<sub>2</sub>(t+2)との間で和音の入れ替えを行っても良い。また、M<sub>1</sub>(t-2)=M<sub>1</sub>(t+1), M<sub>2</sub>(t-2)=M<sub>2</sub>(t+1), M<sub>1</sub>(t-1)=M<sub>1</sub>(t)=M<sub>1</sub>(t+1)=M<sub>2</sub>(t-2)及びM<sub>2</sub>(t-1)=M<sub>2</sub>(t)=M<sub>2</sub>(t+1)=M<sub>1</sub>(t-2)の関係が成立するか否かが判別される。この関係が成立する場合には、M<sub>1</sub>(t-1)=M<sub>1</sub>(t)=M<sub>1</sub>(t-2)及びM<sub>2</sub>(t-1)=M<sub>2</sub>(t)=M<sub>2</sub>(t-2)が定められ、M<sub>1</sub>(t-2)とM<sub>2</sub>(t-2)との間で和音の入れ替えが行われる。なお、M<sub>1</sub>(t-2)とM<sub>2</sub>(t-2)との間で和音の入れ替えに代えてM<sub>1</sub>(t+1)とM<sub>2</sub>(t+1)との間で和音の入れ替えを行っても良い。

10

## 【0035】

ステップS41において読み出された第1和音候補M<sub>1</sub>(0)～M<sub>1</sub>(R)及び第2和音候補M<sub>2</sub>(0)～M<sub>2</sub>(R)の各和音が、例えば、図11に示すように時間経過と共に変化する場合には、ステップS42の平均化を行うことによって図12に示すように修正される。更に、ステップS43の和音の入れ替えを行うことによって第1及び第2和音候補の和音の変化は図13に示すように修正される。なお、図11～図13は和音の時間変化を折れ線グラフとして示しており、縦軸は和音の種類に対応した位置となっている。

20

## 【0036】

ステップS43の和音の入れ替え後の第1和音候補M<sub>1</sub>(0)～M<sub>1</sub>(R)のうちの和音が変化した時点tのM<sub>1</sub>(t)及び第2和音候補M<sub>2</sub>(0)～M<sub>2</sub>(R)のうちの和音が変化した時点tのM<sub>2</sub>(t)が各々検出され(ステップS44)、その検出された時点t(4バイト)及び和音(4バイト)が第1及び第2和音候補毎にデータ蓄積装置5に記憶される(ステップS45)。ステップS45で記憶される1楽曲分のデータが和音進行楽曲データである。かかるステップS41～S45が平滑化手段に相当する。

30

## 【0037】

ステップS43の和音の入れ替え後の第1和音候補M<sub>1</sub>(0)～M<sub>1</sub>(R)及び第2和音候補M<sub>2</sub>(0)～M<sub>2</sub>(R)の和音が図14(a)に示すように時間経過と共に変化する場合には、変化時点の時刻と和音とがデータとして抽出される。図14(b)が第1和音候補の変化時点のデータ内容であり、F, G, D, B<sub>b</sub>, Fが和音であり、それらは16進データとして0x08, 0x0A, 0x05, 0x01, 0x08と表される。変化時点tの時刻はT<sub>1</sub>(0), T<sub>1</sub>(1), T<sub>1</sub>(2), T<sub>1</sub>(3), T<sub>1</sub>(4)である。また、図14(c)が第2和音候補の変化時点のデータ内容であり、C, B<sub>b</sub>, F#m, B<sub>b</sub>, Cが和音であり、それらは16進データとして0x03, 0x01, 0x29, 0x01, 0x03と表される。変化時点tの時刻はT<sub>2</sub>(0), T<sub>2</sub>(1), T<sub>2</sub>(2), T<sub>2</sub>(3), T<sub>2</sub>(4)である。図14(b)及び図14(c)に示したデータ内容は楽曲の識別情報と共にデータ蓄積装置5には、ステップS45においては図14(d)に示すような形式で1ファイルとして記憶される。

40

## 【0038】

異なる楽曲音を示すオーディオ信号について上記した和音分析動作を繰り返すことによりデータ蓄積装置5には複数の楽曲毎のファイルとして和音進行楽曲データが蓄積されることになる。なお、データ蓄積装置4にはデータ蓄積装置5の和音進行楽曲データに対応したPCM信号からなる楽曲データが蓄積される。

50

ステップS44において第1和音候補のうちの和音が変化した時点の第1和音候補及び第2和音候補のうちの和音が変化した時点の第2和音候補が各々検出され、それが最終的な和音進行楽曲データとなるので、MP3のような圧縮データに比べても1楽曲当たりの容量を小さくすることができ、また、各楽曲のデータを高速処理することができる。

## 【0039】

また、データ蓄積装置5に書き込まれた和音進行楽曲データは、実際の楽曲と時間的に同期した和音データとなるので、第1和音候補のみ、或いは第1和音候補と第2和音候補との論理和出力を用いて実際に和音を楽曲再生装置10によって生成すれば、楽曲の伴奏が可能となる。

次に、データ蓄積装置5に和音進行楽曲データとして蓄積された楽曲の構造を検出する楽曲構造検出動作について説明する。楽曲構造検出動作は和音進行比較装置7及び繰り返し構造検出装置8によって実行される。

## 【0040】

楽曲構造検出動作においては、図15に示すように、楽曲構造検出対象の楽曲の第1和音候補M1(0)～M1(a-1)及び第2和音候補M2(0)～M2(b-1)が蓄積手段であるデータ蓄積装置5から読み出される(ステップS51)。その楽曲構造検出対象の楽曲は例え、操作入力装置2の操作によって指定される。aは第1和音候補の総数であり、bは第2和音候補の総数である。また、仮想データとして各々K個の第1和音候補M1(a)～M1(a+K-1)及び第2和音候補M2(b)～M2(b+K-1)が用意される(ステップS52)。ここで、 $a < b$ のとき仮想データの第1及び第2和音候補各々の和音総数Pはaに等しく、 $a \geq b$ のとき和音総数Pはbに等しい。仮想データは第1和音候補M1(0)～M1(a-1)及び第2和音候補M2(0)～M2(b-1)の後に付加される。

## 【0041】

読み出された第1和音候補M1(0)～M1(P-1)に対して第1和音差分値MR1(0)～MR1(P-2)が計算される(ステップS53)。第1和音差分値は、 $MR1(0) = M1(1) - M1(0)$ ,  $MR1(1) = M1(2) - M1(1)$ , ……,  $MR1(P-2) = M1(P-1) - M1(P-2)$ の如く計算される。この計算では第1和音差分値MR1(0)～MR1(P-2)各々が0より小であるか否かを判別し、0より小の第1和音差分値には12を加算することが行われる。また、第1和音差分値MR1(0)～MR1(P-2)各々には和音変化後の和音属性MA1(0)～MA1(P-2)が付加される。読み出された第2和音候補M2(0)～M2(P-1)に対しても第2和音差分値MR2(0)～MR2(P-2)が計算される(ステップS54)。第2和音差分値は、 $MR2(0) = M2(1) - M2(0)$ ,  $MR2(1) = M2(2) - M2(1)$ , ……,  $MR2(P-2) = M2(P-1) - M2(P-2)$ の如く計算される。この計算においても第2和音差分値MR2(0)～MR2(P-2)各々が0より小であるか否かを判別し、0より小の第2和音差分値には12を加算することが行われる。また、第2和音差分値MR2(0)～MR2(P-2)各々には和音変化後の和音属性MA2(0)～MA2(P-2)が付加される。なお、和音属性MA1(0)～MA1(P-2), MA2(0)～MA2(P-2)には図9(b)に示した数値が用いられる。

## 【0042】

図16はステップS53及びS54の動作例を説明している。すなわち、和音候補がAm7, Dm, C, F, Em, F, B♭♯の列である場合に、和音差分値は5, 10, 5, 11, 1, 5となり、和音変化後の和音属性は0x02, 0x00, 0x00, 0x02, 0x00, 0x00となる。なお、和音変化後の和音属性がセブンスの場合にはそれに代えてメジャーとしている。セブンスを用いてもその比較演算結果への影響が小さいので、演算量を削減するためである。

## 【0043】

ステップS54の実行後、カウンタ値cが0に初期化される(ステップS55)。そして、第1和音候補M1(0)～M1(P-1)及び第2和音候補M2(0)～M2(P-1)各々のうちのc番目からK個(例え、20)の和音候補(部分楽曲データ)が抽出される(ステップS56)。すなわち、第1和音候補M1(c)～M1(c+K-1)及び第2和音候補M2(c)～M2(c+K-1)が抽出される。 $M1(c) \sim M1(c+K-1) = U1(0) \sim U1(K-1)$ とし、 $M2(c) \sim M2(c+K-1) = U2(0) \sim U2(K-1)$ とする。図17は処理対象の和音進行データのM1(0)～M1(P-1), M2(0)～M2(P-1)及び仮想データに対するU1(0)～U1(K-1), U2(0)～U2(K-1)の関係を示している。

## 【0044】

10

20

30

40

50

ステップ S 5 6 の実行後、部分楽曲データの第 1 和音候補  $U_1(0) \sim U_1(K-1)$  に対して第 1 和音差分値  $U_R 1(0) \sim U_R 1(K-2)$  が計算される（ステップ S 5 7）。ステップ S 5 7 の第 1 和音差分値は、 $U_R 1(0) = U_1(1) - U_1(0)$ ,  $U_R 1(1) = U_1(2) - U_1(1)$ , …,  $U_R 1(K-2) = U_1(K-1) - U_1(K-2)$  の如く計算される。この計算では第 1 和音差分値  $U_R 1(0) \sim U_R 1(K-2)$  各々が 0 より小であるか否かを判別し、0 より小の第 1 和音差分値には 1/2 を加算することが行われる。また、第 1 和音差分値  $U_R 1(0) \sim U_R 1(K-2)$  各々には和音変化後の和音属性  $U_A 1(0) \sim U_A 1(K-2)$  が付加される。また、部分楽曲データの第 2 和音候補  $U_2(0) \sim U_2(K-1)$  に対しても第 2 和音差分値  $U_R 2(0) \sim U_R 2(K-2)$  が計算される（ステップ S 5 8）。第 2 和音差分値は、 $U_R 2(0) = U_2(1) - U_2(0)$ ,  $U_R 2(1) = U_2(2) - U_2(1)$ , …,  $U_R 2(K-2) = U_2(K-1) - U_2(K-2)$  の如く計算される。この計算においても第 2 和音差分値  $U_R 2(0) \sim U_R 2(K-2)$  各々が 0 より小であるか否かを判別し、0 より小の第 2 和音差分値には 1/2 を加算することが行われる。また、第 2 和音差分値  $U_R 2(0) \sim U_R 2(K-2)$  各々には和音変化後の和音属性  $U_A 2(0) \sim U_A 2(K-2)$  が付加される。

10

## 【0045】

ステップ S 5 3 にて得られた第 1 和音差分値  $M_R 1(0) \sim M_R 1(K-2)$  及び和音属性  $M_A 1(0) \sim M_A 1(K-2)$  と、ステップ S 5 7 にて得られた c 番目から K 個の第 1 和音候補  $U_R 1(0) \sim U_R 1(K-2)$  及び和音属性  $U_A 1(0) \sim U_A 1(K-2)$  と、ステップ S 5 8 にて得られた c 番目から K 個の第 2 和音候補  $U_R 2(0) \sim U_R 2(K-2)$  及び和音属性  $U_A 2(0) \sim U_A 2(K-2)$  とに応じて相互相關演算が行われる（ステップ S 5 9）。相互相關演算では相関係数  $COR(t)$  が次式(3)の如く算出される。相関係数  $COR(t)$  が小さいほど類似性が高いことを示す。

20

$$\begin{aligned} COR(t) = & \Sigma 10(|M_R 1(t+k) - U_R 1(k')| + |M_A 1(t+k) - U_A 1(k')| \\ & + |WM 1(t+k+1)/WM 1(t+k) - WU 1(k'+1)/WU 1(k')|) \\ & + \Sigma 10(|M_R 1(t+k) - U_R 2(k')| + |M_A 1(t+k) - U_A 2(k')| \\ & + |WM 1(t+k+1)/WM 1(t+k) - WU 2(k'+1)/WU 2(k')|) \quad \dots \dots (3) \end{aligned}$$

ただし、 $W_U 1()$ ,  $W_M 1()$ ,  $W_U 2()$  は各和音が維持される時間幅、 $t = 0 \sim P - 1$ 、  
 $\Sigma$  演算は  $k = 0 \sim K - 2$  及び  $k' = 0 \sim K - 2$  である。

30

## 【0046】

ステップ S 5 9 の相関係数  $COR(t)$  は  $t$  が  $0 \sim P - 1$  の範囲で各々算出される。また、ステップ S 5 9 の相関係数  $COR(t)$  の演算では飛び越し処理が行われる。飛び越し処理においては、 $(M_R 1(t+k+k1) - U_R 1(k'+k2))$  又は  $(M_R 1(t+k+k1) - U_R 2(k'+k2))$  の最小値が検出される。 $k_1$  及び  $k_2$  各々は  $0 \sim 2$  までのいずれかの整数である。すなわち、 $k_1$  及び  $k_2$  各々を  $0 \sim 2$  までの範囲で変化させて  $(M_R 1(t+k+k1) - U_R 1(k'+k2))$  又は  $(M_R 1(t+k+k1) - U_R 2(k'+k2))$  の最小値となるときが検出される。そのときの  $k+k_1$  が新たな  $k$  に、 $k'+k_2$  が新たな  $k'$  とされる。その後、式(3)に応じて相関係数  $COR(t)$  が算出される。

40

## 【0047】

更に、各時点の和音から変化後の和音が処理対象の和音進行楽曲データ及びその和音進行楽曲データの c 番目から K 個の部分楽曲データが C 及び A\_m のいずれであっても、或いは C\_m 及び E\_b のいずれであっても同一とみなす。すなわち、変化後の和音が関係調の和音であれば、上記の式の  $|M_R 1(t+k) - U_R 1(k')| + |M_A 1(t+k) - U_A 1(k')| = 0$  又は  $|M_R 1(t+k) - U_R 2(k')| + |M_A 1(t+k) - U_A 2(k')| = 0$  である。例えば、和音 F から一方のデータが 7 度差でメジャーに変化し、他方のデータが 4 度差でマイナーに変化した場合には同一とし、また和音 F から一方のデータが 7 度差でマイナーに変化し、他方のデータが 10 度差でメジャーに変化した場合にも同一として処理される。

## 【0048】

更に、ステップ S 5 4 にて得られた第 2 和音差分値  $M_R 2(0) \sim M_R 2(K-2)$  及び和音属性

50

MA 2(0)～MA 2(K-2)と、ステップS 5 7にて得られたc番目からK個の第1和音候補UR 1(0)～UR 1(K-2)及び和音属性UA 1(0)～UA 1(K-2)と、ステップS 5 8にて得られたc番目からK個の第2和音候補UR 2(0)～UR 2(K-2)及び和音属性UA 2(0)～UA 2(K-2)とに応じて相互相関演算が行われる(ステップS 6 0)。相互相関演算では相関係数COR'(t)が次式(4)の如く算出される。相関係数COR'(t)が小さいほど類似性が高いことを示す。

$$\begin{aligned} \text{COR}'(t) = & \sum 10(|\text{MR } 2(t+k)-\text{UR } 1(k')| + |\text{MA } 2(t+k)-\text{UA } 1(k')| \\ & + |\text{WM } 2(t+k+1)/\text{WM } 2(t+k)-\text{WU } 1(k'+1)/\text{WU } 1(k')|) \\ & + \sum 10(|\text{MR } 2(t+k)-\text{UR } 2(k')| + |\text{MA } 2(t+k)-\text{UA } 2(k')| \\ & + |\text{WM } 2(t+k+1)/\text{WM } 2(t+k)-\text{WU } 2(k'+1)/\text{WU } 2(k')|) \quad \dots\dots(4) \end{aligned}$$

10

ただし、W U 1(), W M 2(), W U 2()は各和音が維持される時間幅、 $t = 0 \sim P - 1$ 、 $\Sigma$ 演算は $k = 0 \sim K - 2$ 及び $k' = 0 \sim K - 2$ である。

#### 【0049】

ステップS 6 0の相関係数COR'(t)は $t$ が $0 \sim P - 1$ の範囲で各々算出される。また、ステップS 6 0の相関係数COR(t)の演算では上記のステップS 5 9と同様に飛び越し処理が行われる。飛び越し処理においては、( $\text{MR } 2(t+k+k_1)-\text{UR } 1(k'+k_2)$ )又は( $\text{MR } 2(t+k+k_1)-\text{UR } 2(k'+k_2)$ )の最小値が検出される。 $k_1$ 及び $k_2$ 各々は $0 \sim 2$ までのいずれかの整数である。すなわち、 $k_1$ 及び $k_2$ 各々を $0 \sim 2$ までの範囲で変化させて( $\text{MR } 2(t+k+k_1)-\text{UR } 1(k'+k_2)$ )又は( $\text{MR } 2(t+k+k_1)-\text{UR } 2(k'+k_2)$ )の最小値となるときが検出される。そのときの $k+k_1$ が新たな $k$ に、 $k'+k_2$ が新たな $k'$ とされる。その後、式(4)に応じて相関係数COR'(t)が算出される。

20

#### 【0050】

更に、各時点の和音から変化後の和音が処理対象の和音進行楽曲データ及び部分楽曲データがC及びAmのいずれであっても、或いはCm及びEbのいずれであっても同一とみなす。すなわち、変化後の和音が関係調の和音であれば、上記の式の $|\text{MR } 2(t+k)-\text{UR } 1(k')| + |\text{MA } 2(t+k)-\text{UA } 1(k')| = 0$ 又は $|\text{MR } 2(t+k)-\text{UR } 2(k')| + |\text{MA } 2(t+k)-\text{UA } 2(k')| = 0$ である。

30

#### 【0051】

図18(a)は処理対象の和音進行楽曲データとその部分楽曲データとの関係を示している。部分楽曲データは $t$ の進行に従って処理対象の和音進行楽曲データとの比較部分が変化する。図18(b)は相関係数COR(t)又はCOR'(t)の変化を示している。ピーク波形部分が類似性が高い部分である。

図18(c)は処理対象の和音進行楽曲データとその部分楽曲データとの相互相関演算における、各和音が維持される時間幅W U(1)～W U(5)、飛び越し処理部分及び関係調の部分を示している。処理対象の和音進行楽曲データと部分楽曲データとの間の矢印線は同一和音を示している。その矢印線のうちの同一時間にない傾いた矢印線で結ばれた和音は、飛び越し処理で検出された和音である。また、矢印線が波線になっているものは関係調の和音である。

40

#### 【0052】

ステップS 5 9及びS 6 0で算出された相関係数COR(t)及びCOR'(t)は加算されて合計相関係数COR(c,t)が算出される(ステップS 6 1)。すなわち、COR(c,t)は次式(5)に示すように算出される。

$$\text{COR}(c, t) = \text{COR}(t) + \text{COR}'(t) \quad t = 0 \sim P - 1 \quad \dots\dots(5)$$

図19(a)～(f)は処理対象の和音進行楽曲データが示す楽曲中のフレーズ(和音進行列)と、部分楽曲データが示すフレーズと、合計相関係数COR(c,t)との関係を示している。和音進行楽曲データが示す楽曲中のフレーズは図示しないイントロIの後の曲の流れ順にA, B, C, A', C', D, C"であり、AとA'が同一フレーズ、またCとC'、C"が同一フレーズとする。図19(a)では、部分楽曲データの先頭にフレーズAが位置

50

している場合にであり、 $COR(c, t)$ は和音進行楽曲データのフレーズAとA'に対応した時点で□で示すピーク値を生成する。図19(b)では、部分楽曲データの先頭にフレーズBが位置している場合にであり、 $COR(c, t)$ は和音進行楽曲データのフレーズBだけに対応した時点で×で示すピーク値を生成する。図19(c)では、部分楽曲データの先頭にフレーズCが位置している場合にであり、 $COR(c, t)$ は和音進行楽曲データのフレーズC, C', C''の各々に対応した時点で○で示すピーク値を生成する。図19(d)では、部分楽曲データの先頭にフレーズA'が位置している場合にであり、 $COR(c, t)$ は和音進行楽曲データのフレーズA, A'の各々に対応した時点で□で示すピーク値を生成する。図19(e)では、部分楽曲データの先頭にフレーズC'が位置している場合にであり、 $COR(c, t)$ は和音進行楽曲データのフレーズC, C', C''の各々に対応した時点で○で示すピーク値を生成する。図19(f)では、部分楽曲データの先頭にフレーズC''が位置している場合にであり、 $COR(c, t)$ は和音進行楽曲データのフレーズC, C', C''の各々に対応した時点で○で示すピーク値を生成する。

10

## 【0053】

ステップS61の実行後、カウンタ値cに1が加算され(ステップS62)、そのカウンタ値cがP-1より大であるか否かが判別される(ステップS63)。 $c \leq P-1$ であるならば、処理対象の和音進行楽曲データ全てに亘って相関係数 $COR(c, t)$ が算出されていない。よって、ステップS56に戻って上記のステップS56～S63の動作が繰り返される。

20

## 【0054】

$c > P-1$ であるならば、 $COR(c, t)$ 、すなわち $COR(0, 0) \sim COR(P-1, P-1)$ のピーク値が検出され、そのピーク値の検出時のc, tについて $COR\_PEAK(c, t) = 1$ が設定され、ピーク値でないときのc, tについて $COR\_PEAK(c, t) = 0$ が設定される(ステップS64)。 $COR(c, t)$ が所定値を越えた部分の最高値をピーク値とする。ステップ64によって $COR\_PEAK(c, t)$ の列が形成される。次に、この $COR\_PEAK(c, t)$ 列において、tが $0 \sim P-1$ の各々の $COR\_PEAK(c, t)$ の合計値がピーク数 $PK(t)$ として算出される(ステップS65)。 $PK(0) = COR\_PEAK(0, 0) + COR\_PEAK(1, 0) + \dots + COR\_PEAK(P-1, 0)$ ,  $PK(1) = COR\_PEAK(0, 1) + COR\_PEAK(1, 1) + \dots + COR\_PEAK(P-1, 1)$ , ...,  $PK(P-1) = COR\_PEAK(0, P-1) + COR\_PEAK(1, P-1) + \dots + COR\_PEAK(P-1, P-1)$ である。ピーク数 $PK(0) \sim PK(P-1)$ のうちの連続する2以上の同一数の範囲が同一フレーズ範囲として区分けされ、それに基づいて楽曲構造データがデータ蓄積装置5に保存される(ステップS66)。例えば、ピーク数 $PK(t)$ が2である場合には、楽曲中で2回繰り返しが行われるフレーズとなり、ピーク数 $PK(t)$ が3である場合には、楽曲中で3回繰り返しが行われるフレーズとなる。同一フレーズの範囲のピーク数 $PK(t)$ は同一値となる。ピーク数 $PK(t)$ が1である場合には、繰り返しがないフレーズを示すことになる。

30

## 【0055】

図20は図19(a)～(f)に示したフレーズI, A, B, C, A', C', D, C''を有する楽曲についてのピーク数 $PK(t)$ と、相関係数 $COR(c, t)$ の算出結果からピーク値が得られた位置 $COR\_PEAK(c, t)$ とを示している。 $COR\_PEAK(c, t)$ はマトリックスで表示しており、横軸が和音数 $t = 0 \sim P-1$ であり、縦軸が部分楽曲データの開始位置である $c = 0 \sim P-1$ を示している。ドット部分が $COR(c, t)$ がピーク値を得た $COR\_PEAK(c, t) = 1$ に対応した位置である。対角線上は同一データ同士の自己相関をとったことになるので、ドット列となる。対角線以外の部分に現れるドット列が繰り返しの和音進行によるフレーズに対応する。図19(a)～(f)に対応して×は1回だけのフレーズI, B, Dに対応し、○は3回の繰り返しフレーズC, C', C''に対応し、□は2回の繰り返しフレーズA, A'に対応する。ピーク数 $PK(t)$ はフレーズI, A, B, C, A', C', D, C''に対応して1, 2, 1, 3, 2, 3, 1, 3となる。これが結果として楽曲構造を示すことになる。

40

## 【0056】

50

楽曲構造データは図21に示すようなフォーマットを有している。各フレーズの開始時刻情報及び終了時刻情報には図14(c)に示した和音進行楽曲データT(t)が用いられる。

また、楽曲構造検出結果が表示装置9に表示される(ステップS67)。楽曲構造検出結果の表示画面は図22に示すように楽曲中の各繰り返しフレーズ部分の選択ができるようになっている。この表示画面によって選択された繰り返しフレーズ部分又は繰り返し回数が最も多いフレーズ部分に対応する楽曲データがデータ蓄積装置4から読み出されて楽曲再生装置10に供給される(ステップS68)。これにより、楽曲再生装置10は供給された楽曲データを順次再生し、それがデジタル信号としてデジタル/アナログ変換装置11に供給される。デジタル/アナログ変換装置11においてアナログオーディオ信号に変換された後、スピーカ15から繰り返しフレーズ部分の再生音が出力されることになる。10

#### 【0057】

よって、利用者は処理対象の楽曲の構造を表示画面から知ることができると共に、その楽曲のうちの選択した繰り返しフレーズ部分又は繰り返し回数が最も多いフレーズ部分を容易に聴取することができる。

上記の楽曲構造検出動作のステップS56が部分楽曲データ生成手段に対応し、ステップS57～S63が類似度(相関係数COR(C, t))を算出する比較手段に相当し、ステップS64が和音位置検出手段に相当し、ステップS65～S68が出力手段に相当する。

#### 【0058】

上記した飛び越し処理及び関係調処理は、和音の変化前後の差分値の演算の際に処理対象の和音進行楽曲データがアナログ信号に基づいて作成された場合における外部雑音や入力装置の周波数特性の影響を排除するため、或いは1番と2番とでは同一フレーズであってもリズムや旋律の変化があったり、又は転調が行われている場合にはデータ間の和音の位置や属性が完全に一致しないことが起きるので、それを防止するために行われる。すなわち、一時的に和音進行が異なっても一定時間幅内で和音進行の傾向が類似していることを検出することができるので、リズムや旋律の変化があったり、又は転調が行われている場合でもその影響を受けることなく、同一フレーズであるか否かを正確に判別することができる。更に、飛び越し処理及び関係調処理を施すことによってその施した部分以外の相互相関演算においても正確な類似度を求めることができる。20

#### 【0059】

また、上記した実施例においては、PCMデータ形式の楽曲データに対して作用することを前提としているが、ステップS28の処理において楽曲に含まれる音符列が分かっていれば、楽曲データとしてMIDIデータを用いることもできる。更に、上記した実施例のシステムを応用すれば、楽曲を構成する繰り返し回数の多いフレーズ部分だけを順に再生する、例えば、ハイライト再生システムを実現することも容易に可能である。

#### 【0060】

図23は本発明の他の実施例を示している。図23の楽曲処理システムにおいては、図1のシステム中の和音解析装置3、一時記憶メモリ6、和音進行比較装置7及び繰り返し構造検出装置8がコンピュータ21によって形成されている。コンピュータ21は記憶装置22に記憶されたプログラムに応じて上記の和音解析動作及び楽曲構造検出動作を実行する。記憶装置22はハードディスクドライブに限らず、記録媒体のドライブ装置でも良い。その記録媒体のドライブ装置の場合には記録媒体に和音進行楽曲データを書き込むようにしても良い。40

#### 【0061】

以上のように、本発明によれば、和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成する部分楽曲データ生成手段と、部分楽曲データ各々と和音進行楽曲データとを和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して複数の楽曲毎の類似度を算出する比較手段と、部分楽曲データ毎に比較手段によって算出された類似度各々に応じて類似50

度が所定値より高いピーク値となった和音進行楽曲データ中の和音の位置を検出する和音位置検出手段と、和音進行楽曲データ中の和音の位置毎に部分楽曲データ全てについて類似度が所定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する出力手段と、を備えたことにより、繰り返し部分を含む楽曲の構造を簡単な構成で適切に検出することができる。

#### 【図面の簡単な説明】

【図1】本発明を適用した楽曲処理システムの構成を示すブロック図である。  
 【図2】周波数誤差検出動作を示すフローチャートである。  
 【図3】A音を1.0とした場合の12音及び1オクターブ高いA音各々の周波数比を示す図である。

10

【図4】和音解析動作の本処理を示すフローチャートである。  
 【図5】帯域データの各音成分の強度レベル例を示す図である。  
 【図6】帯域データの各音成分の強度レベル例を示す図である。  
 【図7】4音からなる和音に対する3音からなる和音への変換を示す図である。  
 【図8】一時記憶メモリへの記録フォーマットを示す図である。  
 【図9】基本音及び和音の属性の表記方法、並びに和音候補の表記方法を示す図である。  
 【図10】和音解析動作の後処理を示すフローチャートである。  
 【図11】平滑化処理前の第1及び第2和音候補の時間変化を示す図である。  
 【図12】平滑化処理後の第1及び第2和音候補の時間変化を示す図である。  
 【図13】入れ替え処理後の第1及び第2和音候補の時間変化を示す図である。  
 【図14】和音進行楽曲データの作成方法及びそのフォーマットを示す図である。  
 【図15】楽曲構造検出動作を示すフローチャートである。

20

【図16】和音変化の和音差分値及び変化後の属性の例を示す図である。  
 【図17】仮想データを含む和音進行楽曲データと部分楽曲データとの関係を示す図である。  
 【図18】相互相関演算時の和音進行楽曲データと部分楽曲データとの関係、相関係数C<sub>OR</sub>(c, t)の変化、並びに各和音が維持される時間幅、飛び越し処理部分及び関係調の部分を示している。

【図19】部分楽曲データに含まれるフレーズと和音進行楽曲データに含まれるフレーズ列とに応じた相関係数C<sub>OR</sub>(c, t)の変化を示す図である。

30

【図20】図19に示したフレーズ列を有する楽曲についてのピーク数PK(t)と、ピーク値が得られた位置C<sub>OR</sub>\_PEAK(c, t)とを示す図である。

【図21】楽曲構造データのフォーマットを示す図である。

【図22】表示装置の表示例を示す図である。

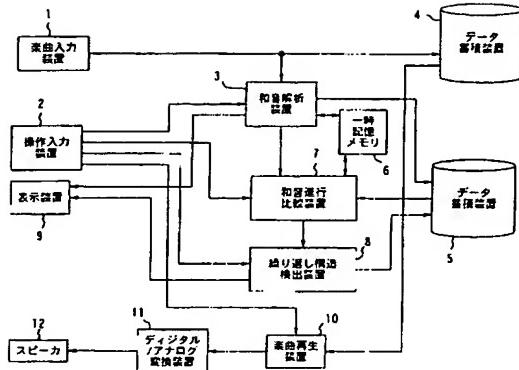
【図23】本発明の他の実施例として楽曲処理システムの構成を示すブロック図である。

#### 【符号の説明】

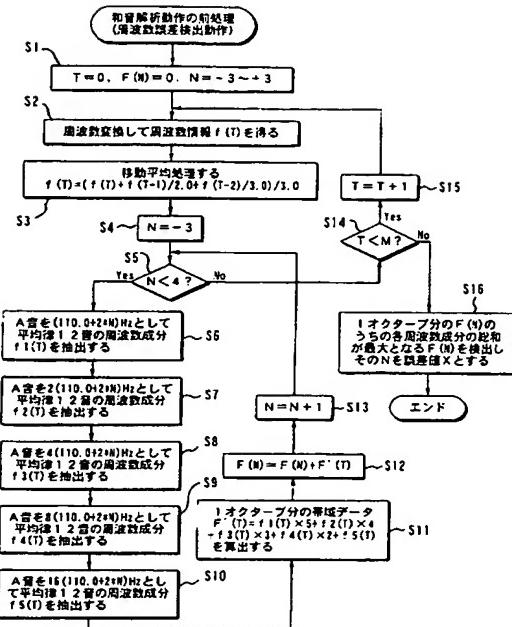
- 3 和音解析装置
- 4, 5 データ蓄積装置
- 7 和音進行比較装置
- 8 繰り返し構造検出装置
- 10 楽曲再生装置
- 21 コンピュータ

40

【図1】



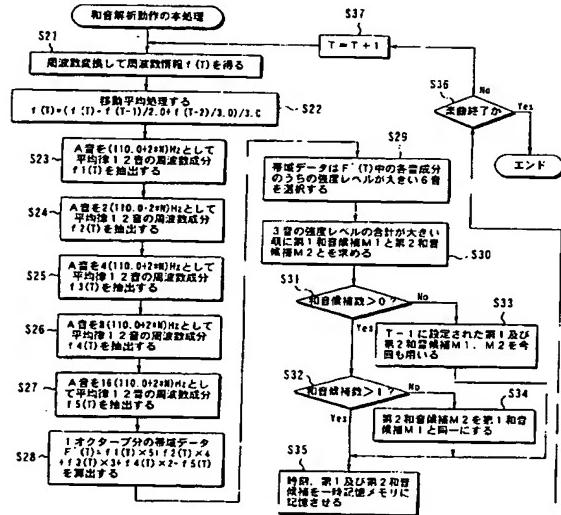
【図2】



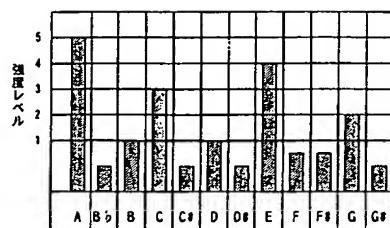
【図3】

A	1.0
A #	AX1.059163
B	AX1.124622
C	AX1.189207
C #	AX1.259921
D	AX1.334840
D #	AX1.414214
E	AX1.498307
F	AX1.587401
F #	AX1.681793
G	AX1.781797
G #	AX1.887749
A	AX2.0

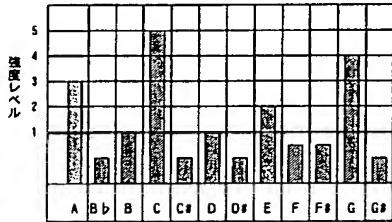
【図4】



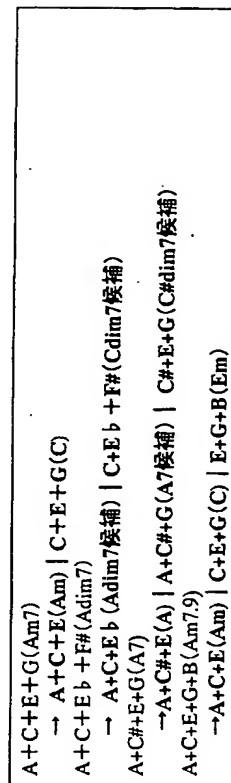
【図5】



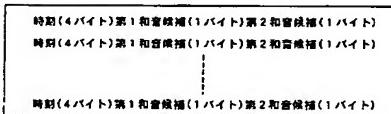
【図6】



【図7】



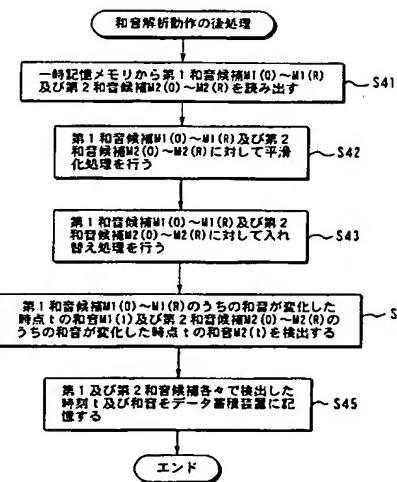
【図8】



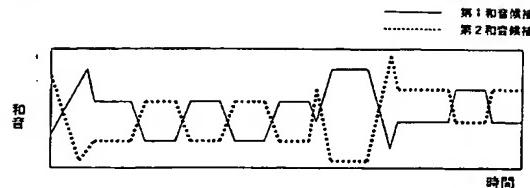
【図9】

属性																																			
(a)	(b)																																		
<table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>基本音</td><td></td></tr> <tr><td>A</td><td>0x00</td></tr> <tr><td>A♯</td><td>0x01</td></tr> <tr><td>B</td><td>0x02</td></tr> <tr><td>C</td><td>0x03</td></tr> <tr><td>C♯</td><td>0x04</td></tr> <tr><td>D</td><td>0x05</td></tr> <tr><td>D♯</td><td>0x06</td></tr> <tr><td>E</td><td>0x07</td></tr> <tr><td>F</td><td>0x08</td></tr> <tr><td>F♯</td><td>0x09</td></tr> <tr><td>G</td><td>0x0A</td></tr> <tr><td>G♯</td><td>0x0B</td></tr> </table>	基本音		A	0x00	A♯	0x01	B	0x02	C	0x03	C♯	0x04	D	0x05	D♯	0x06	E	0x07	F	0x08	F♯	0x09	G	0x0A	G♯	0x0B	<table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>メジャー</td><td>0x00</td></tr> <tr><td>マイナー</td><td>0x02</td></tr> <tr><td>セブンス候補</td><td>0x01</td></tr> <tr><td>Dim7候補</td><td>0x03</td></tr> </table>	メジャー	0x00	マイナー	0x02	セブンス候補	0x01	Dim7候補	0x03
基本音																																			
A	0x00																																		
A♯	0x01																																		
B	0x02																																		
C	0x03																																		
C♯	0x04																																		
D	0x05																																		
D♯	0x06																																		
E	0x07																																		
F	0x08																																		
F♯	0x09																																		
G	0x0A																																		
G♯	0x0B																																		
メジャー	0x00																																		
マイナー	0x02																																		
セブンス候補	0x01																																		
Dim7候補	0x03																																		
属性(下位4ビット)+基本音(下位4ビット)																																			
(c)																																			

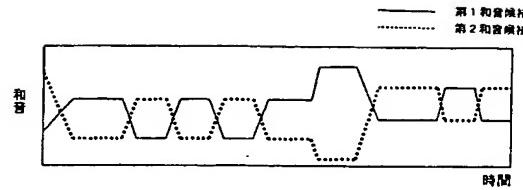
【図10】



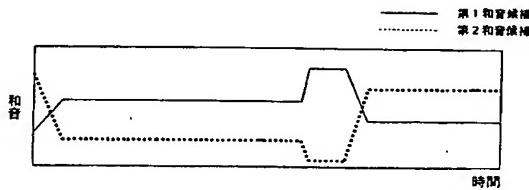
【図11】



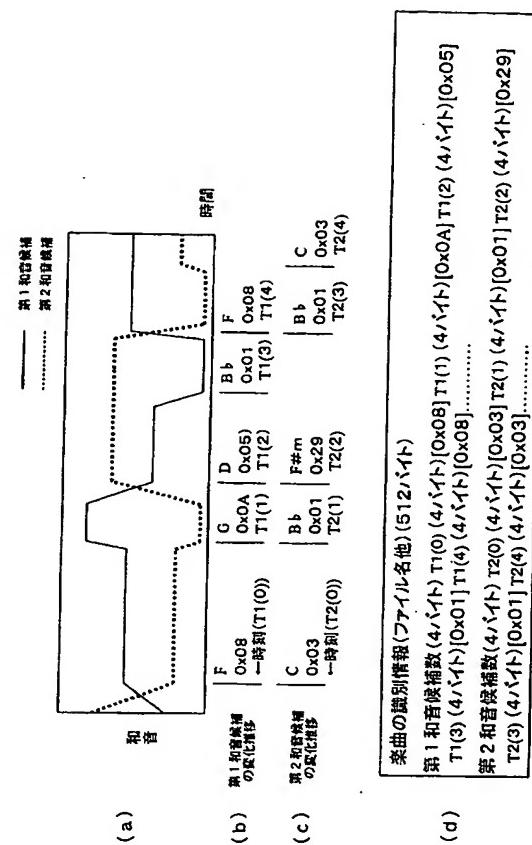
【図12】



【図13】

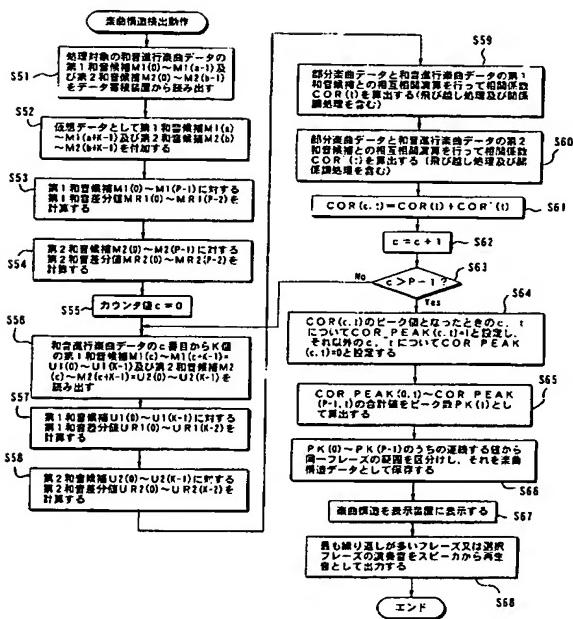


【図14】

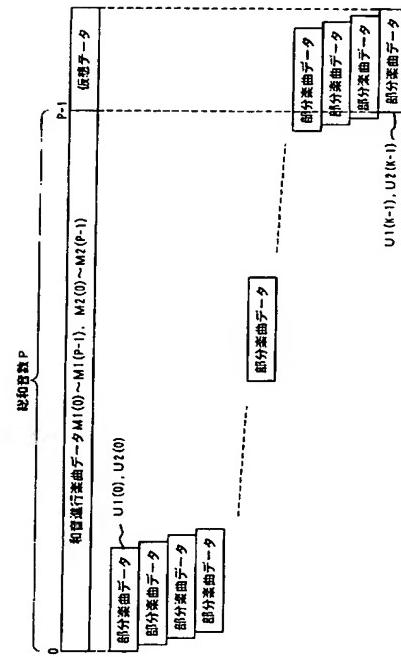


楽曲の箇別情報(ファイル名他) (512バイト)  
 第1和音候補数(4バイト) T1(0) (4バイト)[0x0B] T1(1) (4バイト)[0xA] T1(2) (4バイト)[0x05]  
 T1(3) (4バイト)[0x01] T1(4) (4バイト)[0x0B] .....  
 第2和音候補数(4バイト) T2(0) (4バイト)[0x03] T2(1) (4バイト)[0x01] T2(2) (4バイト)[0x29]  
 T2(3) (4バイト)[0x01] T2(4) (4バイト)[0x03] .....

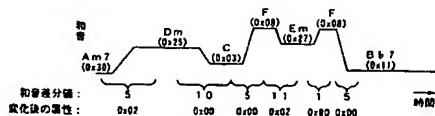
【図15】



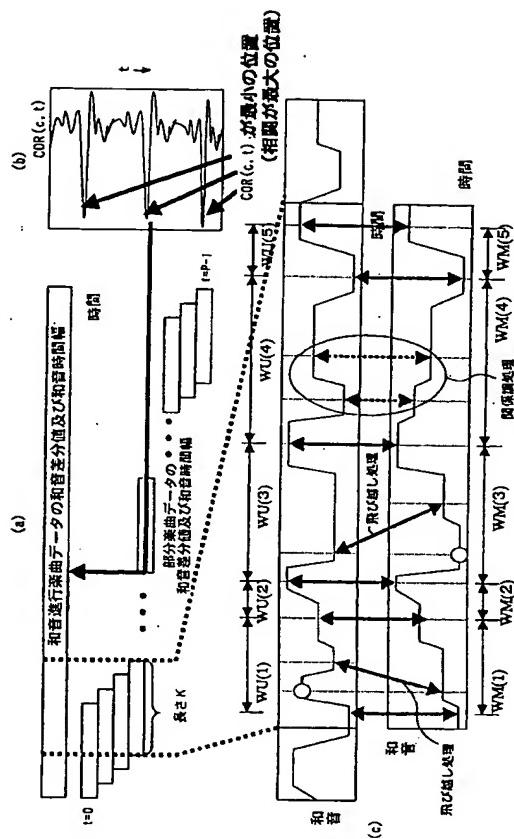
【図17】



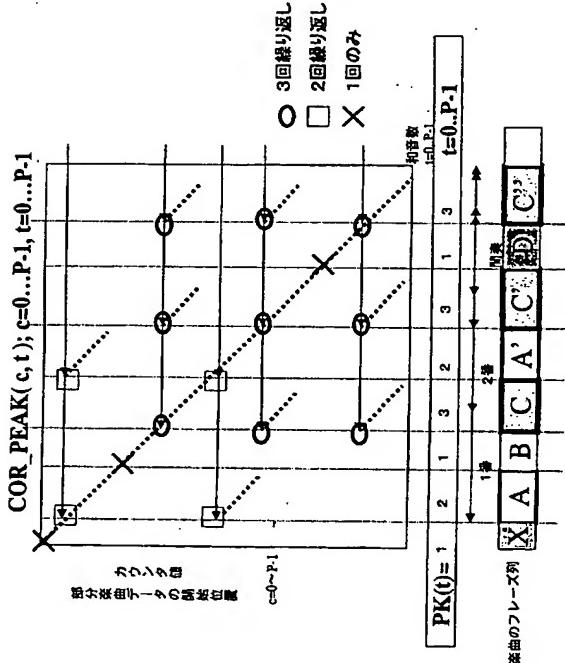
【図16】



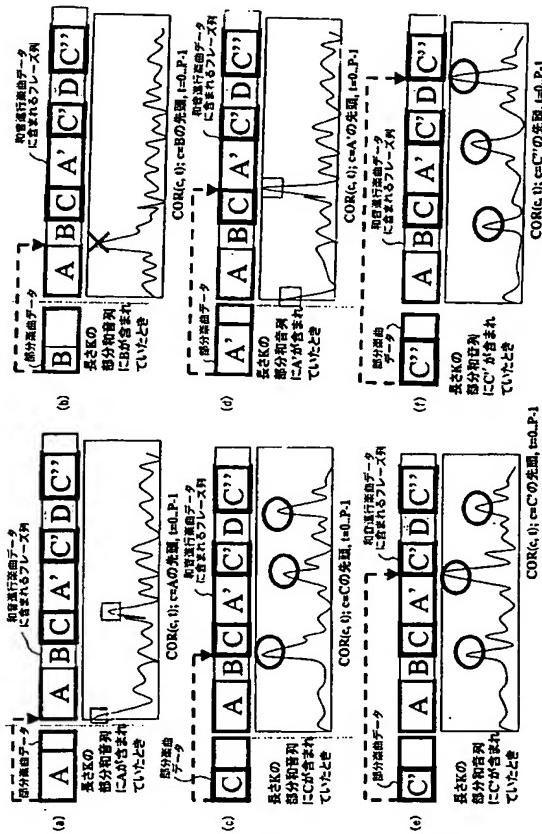
【図18】



【図20】



【図 19】



【図 2-1】

楽曲の概要情報(ファイル名他)(512バイト)  
(繰り返し部分1について: 3回繰り返されるフレーズ)  
0x03(ランク3, 1バイト)第1フレーズ開始時刻情報(4バイト)第1フレーズ終了時刻情報(4バイト)  
第2フレーズ開始時刻情報(4バイト)第2フレーズ終了時刻情報(4バイト)  
第3フレーズ開始時刻情報(4バイト)第3フレーズ終了時刻情報(4バイト)  
(繰り返し部分2について: 2回繰り返されるフレーズ)  
0x02(ランク2, 1バイト)第1フレーズ開始時刻情報(4バイト)第1フレーズ終了時刻情報(4バイト)  
第2フレーズ開始時刻情報(4バイト)第2フレーズ終了時刻情報(4バイト)  
(繰り返し部分3について: 1回繰り返されるフレーズ)  
0x02(ランク2, 1バイト)第1フレーズ開始時刻情報(4バイト)第1フレーズ終了時刻情報(4バイト)  
第2フレーズ開始時刻情報(4バイト)第2フレーズ終了時刻情報(4バイト)  
(繰り返し部分4(または5)について: 1回のみ出現するもの)(1バイト)  
0x01(ランク2, 1バイト)フレーズ開始時刻情報(4バイト)フレーズ終了時刻情報(4バイト)  
0x01(ランク2, 1バイト)フレーズ開始時刻情報(4バイト)フレーズ終了時刻情報(4バイト)  
0x01(ランク2, 1バイト)フレーズ開始時刻情報(4バイト)フレーズ終了時刻情報(4バイト)

【図22】

【図23】

